The structural impact of implementing Optimal Contribution selection in a commercial pig breeding population

The structural impact of implementing Optimal Contribution selection in a commercial pig breeding population

D. M. Howard[1], R. Pong-Wong[1], P. W. Knap[2], V. D. Kremer[2] and J. A. Woolliams[1]
[1]The Roslin Institute and R(D)SVS, University of Edinburgh, Midlothian, United Kingdom
[2]Genus-PIC, Schleswig, Germany

ABSTRACT: This study represents the first analysis of the structural changes occurring in a commercial population following the implementation of Optimal Contribution (OC) selection. It was shown that after OC selection, the estimated Mendelian sampling term ($\hat{a}$) becomes an important determinant for whether an individual made a long-term genetic contribution (r) to the population, as well as for determining the magnitude of that contribution. This provides the first practical validation of published theoretical predictions, in particular that the selective advantage underlying OC is derived from the inclusion of $\hat{a}$ for making selection decisions. This validation is important as genomic evaluation will increase accuracy, in part through greater accuracy of $\hat{a}$, and demonstrates that genomic evaluation combined with OC has the potential to deliver more genetic gain for the same rate of inbreeding.
Keywords: optimal contribution; long-term genetic contribution; estimated breeding value; estimated Mendelian sampling term; porcine

Introduction

The selective advantage derived from applying a dynamic Optimal Contribution (OC) based selection algorithm, such as that described by Meuwissen (1997), has formed the conjecture for a number of simulation studies (Grundy et al. (1998); Avendaño et al. (2004); Sorensen et al. (2005)). The consensus view was that the selective advantage is derived from the estimated Mendelian sampling term ($\hat{a}$) of an individual, rather than the estimated breeding value (*EBV*), which includes components that depend on the *EBV*s of the sire and dam. It is then hypothesized that OC will elicit a shift away from using the family-dependent *EBV* towards a greater importance being placed on $\hat{a}$ for making selection decisions. This will result in effective management of the rate of inbreeding (ΔF), as lower emphasis is placed on specific families, whilst selecting on $\hat{a}$ will ensure the proliferation of advantageous breeding values leading to greater genetic gain.

For the first time, this hypothesis was tested in a commercial population, by evaluating the importance placed on *EBV* versus $\hat{a}$ for determining an individual's long term contribution, r, both before and after the implementation of OC. Firstly, an evaluation was made of the attributes that cause the selection of an individual, which was followed by further analysis to determine whether selected individuals had a zero or non-zero r. Finally a bivariate regression analysis was conducted to assess the importance placed on these variables for determining the ultimate magnitude of r, once an individual had been selected. In the population studied, there had been a sufficient number of descendant generations after the introduction of OC to allow r to approach convergence and so make this analysis feasible.

Materials and Methods

Data. The pedigree and *EBV* records of 115,428 pigs born between 1997 and 2012 within a closed nucleus population were analyzed. An OC selection algorithm based on Newman et al. (2009) and Kinghorn (2011) was implemented within this population, with the transition taking place between 2003 and 2005, becoming fully operational by 2006. The approach defined in Woolliams et al. (1999) was used to calculate the individual r to the 2012 cohort. From early 1999 to mid-2012, a total of 10.4 generations were observed, as defined from the calculated contributions. A total of 17,123 individuals representing two generations of ancestors, born between 1999 and 2001, were defined as the Pre-OC group. The Post-OC group was represented by two generations of ancestors, with 25,043 individuals born between early 2006 and mid-2008. The ΔF, based on the diagonal elements of the numerator relationship matrix, remained relatively constant over these periods. Individual $\hat{a}$ were calculated as the deviation of an individual's *EBV* from the average *EBV* of its parents.

Statistical Analyses. The Pre-OC and Post-OC groups were divided by sex, and then further divided into those selected as parents and the remainder. The selected parents were assessed separately as those with $r \geq 0$, and those with $r > 0$. The number of individuals selected as parents was approximately equal between the Pre-OC and Post-OC groups. To provide additional control of dynamic trend changes over time, the ancestors were split into cohorts of 60-day intervals.

Analyses were then conducted to determine whether *EBV* or $\hat{a}$ were significant factors for whether or not an individual was selected to be a parent in each of the Pre/Post-OC by sex sub-classes within cohorts. These analyses were conducted using GenStat v13 (VSN International) using a logistic regression model for binomial errors, with a linear predictor for ancestor j in cohort i:

$$x_{ij} = c_i + \beta_a \hat{a}_{ij} + \beta_{EBV} EBV_{ij} + \varepsilon_{ij} \quad (1),$$

where x_{ij} is the selection score, i.e., 1 if j in cohort i was selected and 0 otherwise, c_i the effect of the cohort, β_a and β_{EBV} are the regression coefficients and ε_{ij} is the error term.

Further analysis of the determining factors for r was examined for *EBV* and $\hat{a}$ after the introduction of OC, using

the Post-OC groups of ancestors. The following linear model was fitted:

$$y_{ij} = c_i + p_j + \varepsilon_{ij} \quad (2),$$

where p_j is the contribution class (unselected, selected but $r = 0$, selected and $r > 0$) and y_{ij} was either *EBV* or $\hat{a}$.

Bivariate regressions were then conducted to identify significant determinants of variation among the contributions of the selected ancestors. This used simple linear regression models fitted to r_{ij}, the contribution of a selected ancestor j in cohort i:

$$r_{ij} = c_i + \beta_a \hat{a}_{ij} + \beta_{EBV} EBV_{ij} + \varepsilon_{ij} \quad (3)$$

Model (3) was fitted to all parents, irrespective of whether they made a long-term contribution, and then separately to all parents with $r_{ij} > 0$.

Results and Discussion

Attributes influencing initial selection. Under truncation selection, which was practiced prior to the introduction of OC selection in this population, the starting point is the list of candidates ranked by their *EBV*. However, Post-OC it was found that both *EBV* and $\hat{a}$ were influential determinants for the probability of a candidate being selected (Table 1). Whilst the regression coefficients were higher for $\hat{a}$, the *EBV* had lower standard errors, and the variance explained was approximately equal. The coefficients were greater for males than for females, reflecting the greater selection intensity applied.

Table 1. Regression coefficients obtained from logistic regression of selection score on estimated breeding value (*EBV*) and estimated Mendelian sampling term ($\hat{a}$) for Post-OC ancestors. Standard errors are given in parentheses.

	Males	Females
EBV	0.326 (0.064)	0.157 (0.017)
$\hat{a}$	0.678 (0.147)	0.439 (0.041)

Selection differentials. The development of the selection differentials were examined for *EBV* and $\hat{a}$ after the introduction of OC, using the Post-OC groups of ancestors. Table 2 shows progressive gains in both *EBV* and $\hat{a}$ for both males and females following selection. Whilst those selected but ultimately not contributing were superior to unselected individuals, the largest differential emerged over generations between those with $r > 0$ and the other groups. This differential was evident for both *EBV* and $\hat{a}$.

Attributes influencing contribution following initial selection. The Pre-OC and Post-OC periods were examined, both among all those selected as parents (Table 3), and among those with $r > 0$ (Table 4), where the latter excludes those selected as parents but which ultimately provided no genetic material to the 2012 cohort.

Table 2. The superiority of estimated breeding value (*EBV*) and estimated Mendelian sampling term ($\hat{a}$) of selected individuals over unselected individuals among Post-OC ancestors. Values presented are deviations from unselected for those selected but ultimately without a long term contribution (r), $r = 0$ and those selected and contributing, $r > 0$. Standard errors are given in parentheses.

	$r = 0$	$r > 0$
Males		
EBV	0.383 (0.207)	2.635 (0.305)
$\hat{a}$	0.292 (0.068)	0.707 (0.099)
Females		
EBV	0.603 (0.046)	1.631 (0.131)
$\hat{a}$	0.261 (0.019)	0.732 (0.054)

In the Pre-OC period, the r of females once selected, was not associated with either *EBV* or $\hat{a}$ ($P > 0.05$). For those females with $r > 0$, there was a suggestive positive relationship with *EBV*. For males, however, the picture was clearer, with *EBV* being a significant determinant of r, irrespective of conditioning on becoming a parent or $r > 0$. In both cases, individuals with greater *EBV* compared to the *EBV* of their contemporaries were associated with a greater r ($P < 0.05$; 1-sided tests). The regression coefficients for $\hat{a}$ were small and negative.

Table 3. Regression coefficients of long-term contributions on estimated breeding value (*EBV*) and estimated Mendelian sampling term ($\hat{a}$) conditional on being selected as a parent. Standard errors are given in parentheses and values are scaled by 10^4.

	Males	Females
Pre-OC		
EBV	11.7 (5.3)	0.68 (0.60)
$\hat{a}$	-7.1 (9.1)	0.70 (1.25)
Post-OC		
EBV	4.8 (3.2)	0.76 (0.52)
$\hat{a}$	-3.0 (5.7)	3.20 (0.93)

Compared with the Pre-OC period, significant associations were observed for the attributes influencing the r of Post-OC females. The contributions of Post-OC females, once selected, had no significant association with EBV, but were positively associated with aˆ ($P < 0.001$ conditional on selection as a parent, $P < 0.05$ conditional on $r > 0$). However, the relationships of contributions with EBV and aˆ were different for Post-OC males. There was no significant relationship with aˆand conditional on $r > 0$, the EBV was positively associated with the magnitude of the contribution ($P < 0.001$). However, the coefficients for males had large standard errors. The regressions for females were more accurately estimated due to the greater number of female parents compared to males.

Genetic gains from applying Optimum Contributions. Figure 1 shows genetic gain over 5 generations of OC

selection using the Post-OC group defined above as a starting point. The graph shows that the gain obtained was steady. This may be due to various reasons, but it was also consistent with the use of OC. Theoretical results (unpublished) and empirical results from simulation (J. Brisbane, personal communication) both predict that the use of OC will reduce the variance of response.

Table 4. Regression coefficients of long-term contributions on estimated breeding value (*EBV*) and estimated Mendelian sampling term ($\hat{a}$) conditional on $r > 0$. Standard errors are given in parentheses and values are scaled by 10^4.

	Males	Females
Pre-OC		
EBV	13.9 (7.7)	4.79 (3.23)
$\hat{a}$	-1.1 (1.2)	-2.95 (5.70)
Post-OC		
EBV	60.6 (15.0)	-3.90 (5.40)
$\hat{a}$	9.8 (22.4)	16.70 (8.05)

Issues related to validation. Validating the practical use of OC needs both pre-OC and post-OC contributions being close to convergence. This will require 10 generations or more of commercial data, with each phase requiring 5 generations to approach convergence. Over such periods, changes in breeding objectives will occur and introduce issues related to comparability of *EBV* over time. However, the analyses presented have minimized the impact of these by making comparisons within a narrowly defined window of generations. In addition, *EBV* and $\hat{a}$ are dynamic and become more accurate as more information is obtained; so those selected and hence having offspring, will have greater accuracy than those that did not become parents. This problem was addressed by first considering the factors influencing initial selection and then by making further analyses within the selected groups.

Conclusions

The results from this analysis, conducted on a commercial pig population under selection, demonstrate (i) that OC is effective in delivering genetic gain, and (ii) the outcomes are in general agreement with the theoretical predictions and simulated studies. The results show that the use of OC promotes $\hat{a}$ as a selective advantage in selection decisions. This was observed in the Post-OC group of ancestors, both in the influence of $\hat{a}$ in the initial selection of individuals to become parents and in differential contributions once selected. In Post-OC males, it was observed that *EBV* continued to be a strong selective advantage. Avendaño et al. (2004) also found in simulated schemes that both $\hat{a}$ and *EBV* determined *r*, although $\hat{a}$ was the primary advantage. Two possibilities may be advanced for the current results for Post-OC males. First, the generations representing Post-OC were very soon after the introduction of OC and may therefore reflect the continuing re-development of the population structure (e.g., exploiting under-utilized individuals). Second, it may be due to the compromises imposed by practical implementation.

The linking of contribution, r_i, to Mendelian sampling term, a_i, is highly desirable as rates of gain are directly related to $\Sigma r_i a_i$ and not $\Sigma r_i A_i$ (Woolliams and Thompson (1994)), where A_i is the breeding value. This is why OC delivers greater genetic gain for a given ΔF, which itself is directly related to Σr_i^2, and is most efficient at utilizing the variance to generate gain.

This validation of the theory is important because genomic evaluation increases accuracy, predominantly by achieving a greater accuracy in the prediction of the true Mendelian sampling term of an individual. The use of genomic assessment in livestock to more precisely discriminate the true Mendelian sampling term, combined with a dynamic OC algorithm, which can fully utilize this information, is expected to generate greater levels of genetic gain for a predefined rate of inbreeding.

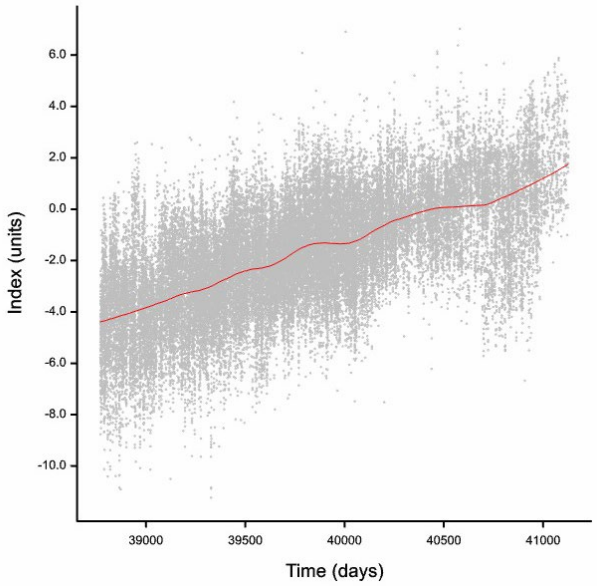

Figure 1. Genetic gain over 5 generations of selection in a commercial pig line following the implementation of OC selection. With the line representing a locally weighted regression fitted to the data with 6 df.

Literature Cited

Avendaño, S., Woolliams, J. A., and Villanueva, B. (2004). Genet. Res. 83:55-64.

Grundy, B., Villanueva, B., and Woolliams, J. A. (1998). Genet. Res. 72:159-168.

Kinghorn, B. P. (2011). Genet. Sel. Evol. 43:4.

Meuwissen, T. H. E. (1997). J. Anim. Sci. 75:934-940.

Newman, S., Kremer, V. D., and Kinghorn, B. P. (2009). 58th An. Nat. Breed. Roundtable. St Louis, Missouri.

Sorensen, A., Berg, P., and Woolliams, J. (2005). Genet. Sel. Evol. 37:57-81.

Woolliams, J. A., and Thompson, R. (1994). Proc. 5th World Cong. Genet. Appl. Livest. Prod. 19:127-134.

Woolliams, J. A., Bijma, P., and Villanueva, B. (1999). Genetics 153:1009-1020.